

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION  
EN MATIÈRE DE BREVETS (PCT)(19) Organisation Mondiale de la Propriété  
Intellectuelle  
Bureau international(43) Date de la publication internationale  
12 février 2004 (12.02.2004)

PCT

(10) Numéro de publication internationale  
WO 2004/014081 A1(51) Classification internationale des brevets<sup>7</sup> : H04N 7/26(21) Numéro de la demande internationale :  
PCT/EP2003/050331

(22) Date de dépôt international : 23 juillet 2003 (23.07.2003)

(25) Langue de dépôt : français

(26) Langue de publication : français

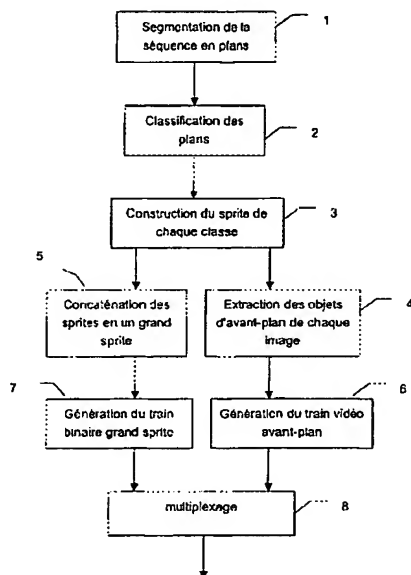
(30) Données relatives à la priorité :  
0209639 30 juillet 2002 (30.07.2002) FR(71) Déposant (pour tous les États désignés sauf US) : THOM-  
SON LICENSING S.A. [FR/FR]; 46 Quai Alphonse Le  
Gallo, F-92100 BOULOGNE-BILLANCOURT (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : FRAN-  
COIS, Edouard [FR/FR]; 18 Allée du Locar, F-35890  
BOURG-DES-COMPTES (FR). THOREAU, Dominique  
[FR/FR]; 39 rue du Réage, F-35510 CESSON-SEVIGNE  
(FR). KYPREOS, Jean [FR/FR]; 15 rue Erik Satie,  
F-35830 BETTON (FR).(74) Mandataire : RUEILLAN-LEMONNIER, Brigitte; 46  
Quai Alphonse Le Gallo, F-92648 BOULOGNE cedex  
(FR).(81) États désignés (national) : AE, AG, AL, AM, AT, AU, AZ,  
BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ,  
DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM,  
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,  
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,

[Suite sur la page suivante]

(54) Title: METHOD FOR COMPRESSING DIGITAL DATA OF A VIDEO SEQUENCE COMPRISING ALTERNATED SHOTS

(54) Titre : PROCÉDE DE COMPRESSION DE DONNÉES NUMÉRIQUES D'UNE SÉQUENCE VIDEO COMPORTANT DES  
PLANS ALTERNES

- 1 ... SEGMENTATION OF THE SEQUENCE INTO SHOTS  
 2 ... CLASSIFICATION OF THE SHOTS  
 3 ... BUILDING OF THE SPRITE OF EACH CLASS  
 4 ... EXTRACTION OF THE FOREGROUND OBJECTS OF EACH IMAGE  
 5 ... CONCATENATION OF THE SPRITES INTO A BIG SPRITE  
 6 ... GENERATION OF THE FOREGROUND VIDEO STREAM  
 7 ... GENERATION OF THE BIG SPRITE BINARY STREAM  
 8 ... MULTIPLEXING

(57) Abstract: The invention provides a method characterised in that it comprises the steps of segmenting (1) a sequence into alternated video shots, classifying (2) said shots according to view-points to obtain classes, building a sprite (3) or video object shot for a class which is an image corresponding to the background relating to said class, merging (5) at least two sprites on a same sprite or video object shot to form an image called big sprite, extracting (4) objects of image foregrounds of the sequence relating to all the shots corresponding to the big sprite, separately coding the big sprite and the extracted foreground objects. Said invention applies to the transmission and storing of video data.

(57) Abrégé : Le procédé est caractérisé en ce qu'il comporte les étapes suivantes: - une segmentation (1) de la séquence en plans alternés vidéo, - une classification (2) de ces plans en fonction de points de vue pour obtenir des classes, - une construction d'un sprite (3) ou plan objet vidéo pour une classe qui est une image correspondant à l'arrière plan relatif à cette classe, - un regroupement (5) d'au moins deux sprites sur un même sprite ou plan objet vidéo pour former une image appelée grand sprite, - une extraction (4), pour les plans correspondant au grand sprite, d'objets d'avant-plan d'images de la séquence relatives à ces plans, - un codage séparé du grand sprite et des objets d'avant-plan extraits. Application à la transmission et au stockage de données vidéo.



MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

- (84) États désignés (*régional*) : brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Publiée :**

- avec rapport de recherche internationale
- avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues

*En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.*

## PROCEDE DE COMPRESSION DE DONNEES NUMERIQUES D'UNE SEQUENCE VIDEO COMPORTANT DES PLANS ALTERNES

L'invention concerne un procédé de compression de données  
5 numériques d'une séquence vidéo composée de plans alternés, à partir de  
"sprites", et un dispositif pour sa mise en œuvre. Elle se situe dans le  
contexte général de la compression vidéo, en particulier dans celui de la  
norme MPEG-4 vidéo.

10 Le terme "sprite" est défini par exemple dans la norme MPEG  
4. Il s'agit d'un objet vidéo (VOP, acronyme de l'anglais Video Object Plane),  
généralement de dimension supérieure à la vidéo affichée, et persistant avec  
le temps. Il est utilisé pour représenter des zones plus ou moins statiques,  
telles que des arrière-plans. Il est codé à partir d'un découpage par  
15 macroblochs. Par la transmission d'un sprite représentant l'arrière-plan  
panoramique et par le codage des paramètres de mouvement décrivant le  
mouvement de la caméra, paramètres représentant par exemple la  
transformée affine du sprite, il est possible de reconstruire des images  
consécutives d'une séquence à partir de ce sprite unique.

20 L'invention concerne en particulier les séquences vidéo  
comprenant une succession de plans générés de façon alternative à partir de  
points de vue similaires. Il peut par exemple s'agir d'une séquence  
d'interview, où l'on voit de façon alternative l'interviewer et l'interviewé,  
25 chacun sur un arrière-plan différent mais en grande partie statique. Cette  
alternance n'est pas limitée à deux points de vue différents. La séquence  
peut être composée de N plans, issus de Q points de vue différents.

Les codages de type classique ne prennent pas en compte ce  
30 type de séquence et le coût de codage ou le taux de compression est donc  
équivalent à celui d'autres séquences. L'approche classique consiste en  
effet, en chaque début de plan, à coder une image en mode intra, à laquelle  
succèdent des images en mode prédictif. Si un plan issu d'un premier point  
de vue apparaît une première fois, suivi d'un plan issu d'un autre point de  
35 vue, suivi d'un plan issu du premier point de vue, la première image de ce  
plan est codée intégralement en mode intra même si une grande partie,

constituée de l'arrière-plan de la scène filmée, est similaire aux images du premier plan. Ceci induit un coût de codage important.

Une solution connue à ce problème de ré-encodage d'un arrière-plan déjà apparu antérieurement consiste à mémoriser, à chaque  
5 détection de changement de plan, la dernière image d'un plan. Au début d'un nouveau plan, la première image est codée par prédiction temporelle ayant pour référence, parmi les images mémorisées, celle qui lui ressemble le plus et qui correspond donc à un même point de vue. Une telle solution peut être considérée comme s'inspirant directement d'un outil connu sous le nom  
10 anglais de "multi-frame referencing", disponible par exemple dans le standard MPEG-4 partie 10 en cours de développement. Une telle solution est cependant consommatrice en mémoire, difficile de mise en œuvre et coûteuse.

15 L'invention a pour but de pallier les inconvénients précités. Elle a pour objet un procédé de compression de données numériques d'une séquence vidéo, caractérisé en ce qu'il comporte les étapes suivantes :

- une segmentation de la séquence en plans alternés vidéo,
- une classification de ces plans en fonction de points de vue  
20 pour obtenir des classes,
  - une construction d'un sprite ou plan objet video pour une classe qui est une image composée correspondant à l'arrière plan relatif à cette classe,
  - un regroupement d'au moins deux sprites sur un même sprite  
25 ou plan objet vidéo pour former une image appelée grand sprite,
  - une extraction, pour les plans correspondant au grand sprite, d'objets d'avant-plan d'images de la séquence relatives à ces plans,
  - un codage séparé du grand sprite et des objets d'avant-plan extraits.

30 Selon une mise en œuvre particulière, les sprites sont placés l'un sous l'autre pour construire le grand sprite.

Selon une mise en œuvre particulière, le positionnement des sprites est calculé en fonction du coût de codage du grand sprite.

Le codage exploité est par exemple le codage MPEG-4, le  
35 grand sprite étant alors codé conformément aux sprites définis dans la norme MPEG-4.

Selon une mise en œuvre particulière, le procédé réalise une opération de multiplexage (8) des données relatives aux objets d'avant-plan extraits et des données relatives au grand sprite pour fournir un flux de données.

5 L'invention concerne également le flux de données comprimées pour le codage d'une séquence d'images selon le procédé précédemment décrit, caractérisé en ce qu'il comporte des données de codage du grand sprite associées à des paramètres de déformation applicables au grand sprite et des données de codage des objets d'avant-plan extraits.

10 L'invention concerne également un codeur pour le codage des données selon le procédé précédemment décrit, caractérisé en ce qu'il comporte un circuit de traitement pour la classification des séquences en plans, la construction d'un sprite pour chaque classe et la composition d'un grand sprite par concaténation de ces sprites, un circuit d'extraction d'objets d'avant-plan d'images de la séquence relatives au grand sprite et un circuit  
15 de codage pour le codage du grand sprite et des objets d'avant-plan extraits.

L'invention concerne également un décodeur pour le décodage de données vidéo d'une séquence vidéo comportant des plans alternés selon le procédé précédemment décrit, caractérisé en ce qu'il comporte un circuit  
20 de décodage de données relatives à un grand sprite et de données relatives à des objets d'avant-plan et un circuit de construction d'images à partir des données décodées.

Le sprite est utilisé pour décrire l'arrière-plan de l'ensemble des  
25 plans vidéo issus d'un même point de vue. Ce sprite est codé une seule fois. Ensuite, pour chaque image de ces plans vidéo, le processus consiste à coder les paramètres de déformation à appliquer au sprite pour reconstruire ce qui est perçu de l'arrière-plan dans l'image. Les objets de premier-plan sont, quant à eux, codés comme des objets vidéo ou VOPs (Video Object  
30 Plan) non rectangulaires. Au décodage, ces VOPs sont composés avec l'image de l'arrière-plan pour obtenir l'image finale. Comme la séquence comporte des plans issus de plusieurs points de vue, plusieurs sprites sont nécessaires. Une mise en œuvre particulière de l'invention consiste à concaténer ces différents sprites en un seul grand sprite qui résume alors les  
35 différents arrière-plans de la séquence vidéo complète.

Grâce à l'invention, le ré-encodage de l'arrière-plan, à chaque réapparition de cet arrière-plan, est évité. Le coût de compression de ce type de séquences vidéo est réduit par rapport à un schéma de codage classique de type MPEG-2 ou H.263.

5

D'autres particularités et avantages apparaîtront clairement dans la description suivante donnée à titre d'exemple non limitatif, et faite en regard des figures annexées qui représentent:

- la figure 1, un organigramme d'un procédé de codage selon l'invention,
- la figure 2, l'intégration d'un sprite dans un grand sprite.
- la figure 3, des blocs d'un sprite en bordure haute et basse d'un grand sprite,
- la figure 4, un bloc courant dans son environnement pour le codage par prédiction DC/AC.

10  
15

La figure 1 représente un organigramme simplifié d'un procédé de codage selon l'invention. Ce procédé se scinde en deux phases principales: une phase d'analyse et une phase de codage.

20

La phase d'analyse comprend une première étape 1 qui est une étape de segmentation de la séquence vidéo en plans. Une deuxième étape 2 réalise une classification des plans selon le point de vue dont ils sont issus. Une classe est définie comme un sous-ensemble de plans issus d'un même point de vue. La troisième étape effectue la construction d'un sprite "résumant" l'arrière-plan visible dans les plans du sous-ensemble, ceci pour chacun des sous-ensembles. Pour chaque image de chaque plan du sous-ensemble, des paramètres de déformation, permettant de reconstruire à partir du sprite ce qui est perçu de l'arrière-plan, sont aussi calculés. Une étape de segmentation d'image 4 effectue une segmentation pour chaque image des différents plans, segmentation dans le but de distinguer l'arrière-plan de l'avant-plan. Cette étape permet d'extraire des objets d'avant-plan de chaque image. L'étape 5 est effectuée parallèlement à l'étape 4 et succède donc à l'étape 3. Elle consiste en une concaténation des différents sprites en un seul grand sprite, avec mise à jour des paramètres de déformation prenant en compte la position de chaque sprite dans le grand sprite.

25  
30  
35

La phase de codage succède à la phase d'analyse. Les étapes 6 et 7 succèdent respectivement aux étapes 4 et 5 et génèrent respectivement un train binaire vidéo codant l'avant-plan et un train binaire vidéo codant le grand sprite. Ces trains binaires sont ensuite multiplexés à l'étape 8 pour fournir le train de codage vidéo.

L'étape 1 de segmentation en plans effectue une découpe de la séquence en plans vidéo en comparant les images successives, par exemple en exploitant un algorithme de détection de changement de plans. L'étape 2 de classification compare les différents plans obtenus, à partir de leur contenu, et regroupe dans une même classe les plans similaires, c'est à dire issus d'un point de vue identique ou proche.

L'étape 4 réalise une extraction des objets d'avant-plan. Des masques binaires successifs sont calculés distinguant, pour chaque image de la séquence vidéo, l'arrière-plan de l'avant-plan. A l'issue de cette étape 4, on dispose donc, pour chaque plan, d'une succession de masques, binaires ou non, indiquant les parties de l'avant-plan et de l'arrière-plan. Dans le cas d'un traitement non binaire, le masque correspond en fait à une carte de transparence.

La concaténation des sprites en un grand sprite effectuée à l'étape 5 peut être réalisée de manière à minimiser le coût de codage de ce grand sprite comme proposé ci-dessous. Les informations de codage sont, entre autres, les informations de texture et les informations de déformation. Ces dernières informations sont par exemple les paramètres de déformation successifs qui sont applicables sur le grand sprite, en fonction du temps, et qui sont mises à jour lors de la génération du grand sprite. Ce sont en effet ces paramètres de transformation qui, appliqués au grand sprite, permettront de construire et mettre à jour les fonds nécessaires aux différents plans. Ces informations de codage sont transmises à l'étape 7 pour permettre la génération du train binaire grand sprite.

Dans notre réalisation, deux trains binaires sont générés, l'un codant le grand sprite et l'autre codant l'ensemble des objets de l'avant-plan regroupés en un seul objet. Ces trains binaires sont ensuite multiplexés à l'étape 8. Dans la norme MPEG-4, un flux élémentaire est généré par objet. Il est donc tout aussi envisageable de transmettre plusieurs flux élémentaires ou de ne pas effectuer de multiplexage avec le flux relatif au grand sprite pour la transmission des données codées.

On remarquera que l'étape 4 d'extraction des objets est en fait très corrélée à l'étape précédente de construction d'un sprite, aussi peut-elle être effectuée simultanément, voire même antérieurement, à la précédente. Egalement, les opérations aux étapes 5 et 7 qui sont décrites en parallèle des opérations aux étapes 4 et 6, peuvent être effectuées successivement ou antérieurement à ces étapes 4 et 6. D'autre part, certaines étapes d'analyse, par exemple celle d'extraction des objets, peuvent être évitées dans le cas où l'on dispose d'une description de contenu de type MPEG-7 du document vidéo à coder.

Comme indiqué précédemment, la concaténation peut se faire en cherchant à minimiser le coût de codage du grand sprite. Cela peut porter sur trois points: la texture, la forme, si elle existe, les paramètres de déformation successifs. Cependant le critère prépondérant est le coût de codage de la texture.

Une méthode de minimisation de ce coût est donnée ci-après dans un mode de réalisation exploitant la norme MPEG-4 et effectuant un assemblage des sprites de manière simple, c'est à dire en les superposant horizontalement, méthode qui s'appuie sur le fonctionnement de l'outil de prédiction spatiale DC/AC de MPEG-4. Dans le cadre de la norme MPEG-4, la prédiction spatiale se fait horizontalement ou verticalement. Elle porte de façon systématique sur le premier coefficient DCT de chaque bloc (mode "DC prediction" en anglais dans la norme) et peut aussi, de manière optionnelle, porter sur les autres coefficients DCT de la première ligne ou première colonne de chaque bloc (mode "AC prediction"). Il s'agit de déterminer la position optimale de concaténation, c'est à dire de rechercher le minimum de coût de codage de la texture par un assemblage de sprites voisins présentant sur leurs bords mutuels une continuité de texture.

Le grand sprite est initialisé par le sprite le plus large. Ensuite, un nouveau grand sprite est calculé intégrant le sprite le plus large parmi les sprites restants, c'est à dire le deuxième sprite le plus large. La figure 2 représente un grand sprite 9 et un deuxième grand sprite 10 à intégrer pour obtenir le nouveau grand sprite, c'est à dire à positionner par rapport au sprite 9.

La figure 3 représente le sprite 10 de forme rectangulaire et plus particulièrement la succession de macroblochs 11 en bordure haute et la succession de macroblochs 12 en bordure basse du sprite. Les macroblochs du



sprite pris en compte sont les macroblochs non vides adjacents de la bordure haute lorsque le sprite est placé sous le grand sprite puis de la bordure basse lorsque le sprite est placé au dessus du grand sprite. Dans le cas où le sprite n'est pas rectangulaire, seuls les macroblochs non vides en bordure haute et basse du rectangle englobant ce sprite sont pris en compte. Les macroblochs vides sont ignorés.

Une transformation cosinus discrète DCT est effectuée sur les macroblochs pris en compte (ou blocs luminance des macroblochs), c'est à dire les macroblochs ou blocs non vides en bordure haute et basse des différents sprites. Les positions haute et basse optimales sont ensuite calculées en minimisant un critère de continuité des textures à la frontière des deux sprites.

Pour une position donnée (X,Y) du sprite 10 à intégrer dans le grand sprite 9 précédemment calculé, position définie par des coordonnées (X,Y), une mesure d'un critère global C(X,Y) est calculée. Les positions (X,Y) sont par exemple les coordonnées du coin inférieur gauche du sprite supérieur à intégrer ou les coordonnées du coin supérieur gauche du sprite inférieur à intégrer, l'origine étant définie à partir d'un point prédéterminé du grand sprite. Les coordonnées (X,Y) sont limitées dans la mesure où l'on n'autorise pas le sprite à déborder du grand sprite.

Pour cette position donnée (X,Y) et pour toutes les positions testées, on va avoir N blocs voisins avec le grand sprite, soit situés au-dessus, soit en-dessous. De ces 2 lignes de blocs voisins, c'est à dire celle appartenant au grand sprite et celle appartenant au sprite à intégrer, on considère la ligne des N blocs du dessous. Pour chaque bloc  $B_k$  de ces N blocs, on détermine d'abord quelle sera la direction probable de la prédiction DC/AC.

La figure 4 représente un bloc courant et les blocs environnants, bloc A à sa gauche, bloc B au dessus de A et bloc C au dessus du bloc courant. Comme le fait un outil de prédiction spatiale DC/AC classique, on détermine les gradients des coefficients DC entre les blocs A et B,  $|DC_A - DC_B|$ , et entre les blocs C et B,  $|DC_C - DC_B|$ . S'il n'y a pas de bloc voisin A, B ou C, le coefficient DC est pris par défaut égal à 1024.

- Si  $|DC_A - DC_B| < |DC_C - DC_B|$ , la prédiction DC/AC s'effectuera probablement dans le sens vertical. On va donc déterminer pour

le bloc courant le résidu de sa première ligne correspondant à la prédiction verticale à partir de la première ligne du bloc du dessus C.

- Si  $|DC_A - DC_B| \geq |DC_C - DC_B|$ , la prédiction DC/AC s'effectuera probablement dans le sens horizontal. On va donc déterminer pour le bloc courant le résidu de sa première colonne correspondant à la prédiction horizontale à partir de la première colonne du bloc de gauche A.

On calcule ensuite l'énergie des coefficients AC résiduels, c'est à dire avec prédiction, de la première ligne ou première colonne, selon la direction de prédiction probable:

$$E_{AC\_pred} = \sum_{i=1}^7 (\Delta AC_i)^2$$

$\Delta AC_i$  correspondant au résidu, c'est à dire à la différence entre les 7 coefficients AC de la première ligne ou première colonne du bloc courant et les 7 coefficients AC de la première ligne ou colonne respectivement du bloc supérieur ou du bloc à gauche du bloc courant.

- On calcule également l'énergie des coefficients AC bruts, c'est-à-dire avant prédiction :

$$E_{AC\_brut} = \sum_{i=1}^7 AC_i^2$$

$AC_i$  correspondant aux 7 coefficients AC de la première ligne ou première colonne du bloc courant.

- On cherche à déterminer la position, pour un bloc courant, qui permet d'avoir la plus faible énergie. L'énergie, pour la partie qui varie en fonction de la position du bloc, dépend de  $\Delta DC$  et éventuellement des  $\Delta AC$  s'il y a prédiction. Elle est égale à:

- lorsqu'il y a prédiction DC/AC, c'est à dire si  $E_{AC\_pred} < E_{AC\_brut}$ ,

$$E(B_k) = \Delta DC^2 + \sum_{i=1}^7 (\Delta AC_i)^2$$

- lorsqu'il n'y a pas prédiction DC/AC, c'est à dire si  $E_{AC\_pred} \geq E_{AC\_brut}$ ,

$$E(B_k) = \Delta DC^2$$

Le calcul est effectué pour chacun des blocs N de la ligne et le critère C, pour une position donnée, est alors égal à:

$$C(X, Y) = \sum_{k=1}^N E(B_k)$$

La position optimale  $(X_{opt}, Y_{opt})$  est celle qui minimise  $C(X, Y)$  sur l'ensemble des positions testées.

Une fois déterminé le sprite à intégrer et sa position dans le grand sprite, les paramètres de déformation du sprite à intégrer sont mis à jour. Pour ce faire, il est ajouté à la composante translationnelle de ses paramètres de déformation, les coordonnées ( $X_{opt}$ ,  $Y_{opt}$ ) du point à partir  
5 duquel le nouveau sprite est intégré dans le grand sprite. Dans le cas d'un modèle affine, on a 6 paramètres de déformation (a,b,c,d,e,f), dont 2, a et b, caractérisent la composante translationnelle ou constante de la déformation. Il faut donc transformer a en  $a+X_{opt}$ , et b en  $b+Y_{opt}$ .

10

Les nouveaux paramètres de déformation sont insérés dans la liste des paramètres de déformation du grand sprite, à l'endroit où temporellement le plan correspondant s'insère dans la séquence vidéo.

15

Une fois la concaténation terminée, on dispose

- d'un grand sprite au lieu de plusieurs sprites
- d'une seule liste de paramètres de déformation, au lieu de plusieurs listes correspondant aux différents plans de la séquence vidéo.

20

Les paramètres de déformation successifs permettent de reconstruire, pour chaque image de la séquence vidéo, ce qui est perçu de l'arrière-plan à partir du grand sprite.

25

Le codage peut être effectué en réalisant une passe de pré-analyse de la séquence vidéo suivie d'une passe de codage s'appuyant sur cette analyse.

30

Dans le cas spécifique de la norme MPEG-4, le codage consiste à générer un train binaire en utilisant l'outil de codage sprite (cf. partie 7.8 du document ISO/IEC JTC 1/SC 29/WG 11 N 2502, p.189 à 195). Le second train binaire se base sur les outils de codage d'objets non rectangulaires, en particulier l'outil de codage de la forme binaire (cf. partie 7.5 du document ISO/IEC JTC 1/SC 29/WG 11 N 2502, p.147 à 158), et éventuellement en plus l'outil de codage de la transparence (« grey shape » en anglais, cf. partie 7.5.4 du document ISO/IEC JTC 1/SC 29/WG 11 N 2502, p.160 à 162) si les masques ne sont pas binaires.

35

L'invention concerne également les flux de données comprimées résultant du codage d'une séquence d'images selon le procédé précédemment décrit. Ce flux comporte des données de codage du grand sprite associées à des paramètres de déformation applicables au grand sprite et des données de codage des objets des avant-plans pour la reconstruction des scènes.

L'invention concerne également les codeurs et décodeurs exploitant un tel procédé. Il s'agit par exemple d'un codeur comportant un circuit de traitement pour la classification des séquence en plans, la construction d'un sprite pour chaque classe et la composition d'un grand sprite par concaténation de ces sprites. Il s'agit aussi d'un décodeur comportant un circuit de construction d'images de plans alternés d'une séquence vidéo à partir du décodage de grands sprites et d'objets d'avant plans.

Les applications de l'invention concernent la transmission et le stockage d'images numériques utilisant des normes de codage vidéo avec exploitation de sprites, en particulier la norme MPEG4.

## REVENDECATIONS

1 Procédé de compression de données numériques d'une séquence vidéo, caractérisé en ce qu'il comporte les étapes suivantes :

- 5                   - une segmentation (1) de la séquence en plans alternés vidéo,  
                  - une classification (2) de ces plans en fonction de points de vue pour obtenir des classes,  
                  - une construction d'un sprite (3) ou plan objet video pour une classe qui est une image composée correspondant à l'arrière plan relatif à  
10 cette classe,  
                  - un regroupement (5) d'au moins deux sprites sur un même sprite ou plan objet vidéo, pour former une image appelée grand sprite,  
                  - une extraction (4), pour les plans correspondant au grand sprite, d'objets d'avant-plan d'images de la séquence relatives à ces plans,  
15                   - un codage séparé du grand sprite et des objets d'avant-plan extraits.

2 Procédé selon la revendication 1, caractérisé en ce que les sprites sont placés l'un sous l'autre (5) pour construire le grand sprite.

20

3 Procédé selon la revendication 2, caractérisé en ce que le positionnement des sprites est calculé en fonction du coût de codage du grand sprite.

25

4 Procédé selon la revendication 1, caractérisé en ce que le grand sprite est un sprite tel que défini et codé dans la norme MPEG4.

5 Procédé selon la revendication 1, caractérisé en ce qu'il réalise une opération de multiplexage (8) des données relatives aux objets  
30 d'avant-plan extraits et des données relatives au grand sprite pour fournir un flux de données.

6 Flux de données comprimées pour le codage d'une séquence d'images selon le procédé de la revendication 1, caractérisé en ce qu'il  
35 comporte des données de codage du grand sprite associées à des

paramètres de déformation applicables au grand sprite et des données de codage des objets d'avant-plan extraits.

- 5                   7 Codeur pour le codage des données selon le procédé de la revendication 1, caractérisé en ce qu'il comporte un circuit de traitement pour la classification des séquence en plans, la construction d'un sprite pour chaque classe et la composition d'un grand sprite par concaténation de ces sprites, un circuit d'extraction d'objets d'avant-plan d'images de la séquence relatives au grand sprite et un circuit de codage pour le codage du grand  
10 sprite et des objets d'avant-plan extraits.

- 8 Décodeur pour le décodage de données vidéo d'une séquence vidéo comportant des plans alternés selon le procédé de la revendication 1, caractérisé en ce qu'il comporte un circuit de décodage de  
15 données relatives à un grand sprite et de données relatives à des objets d'avant-plan et un circuit de construction d'images à partir des données décodées.

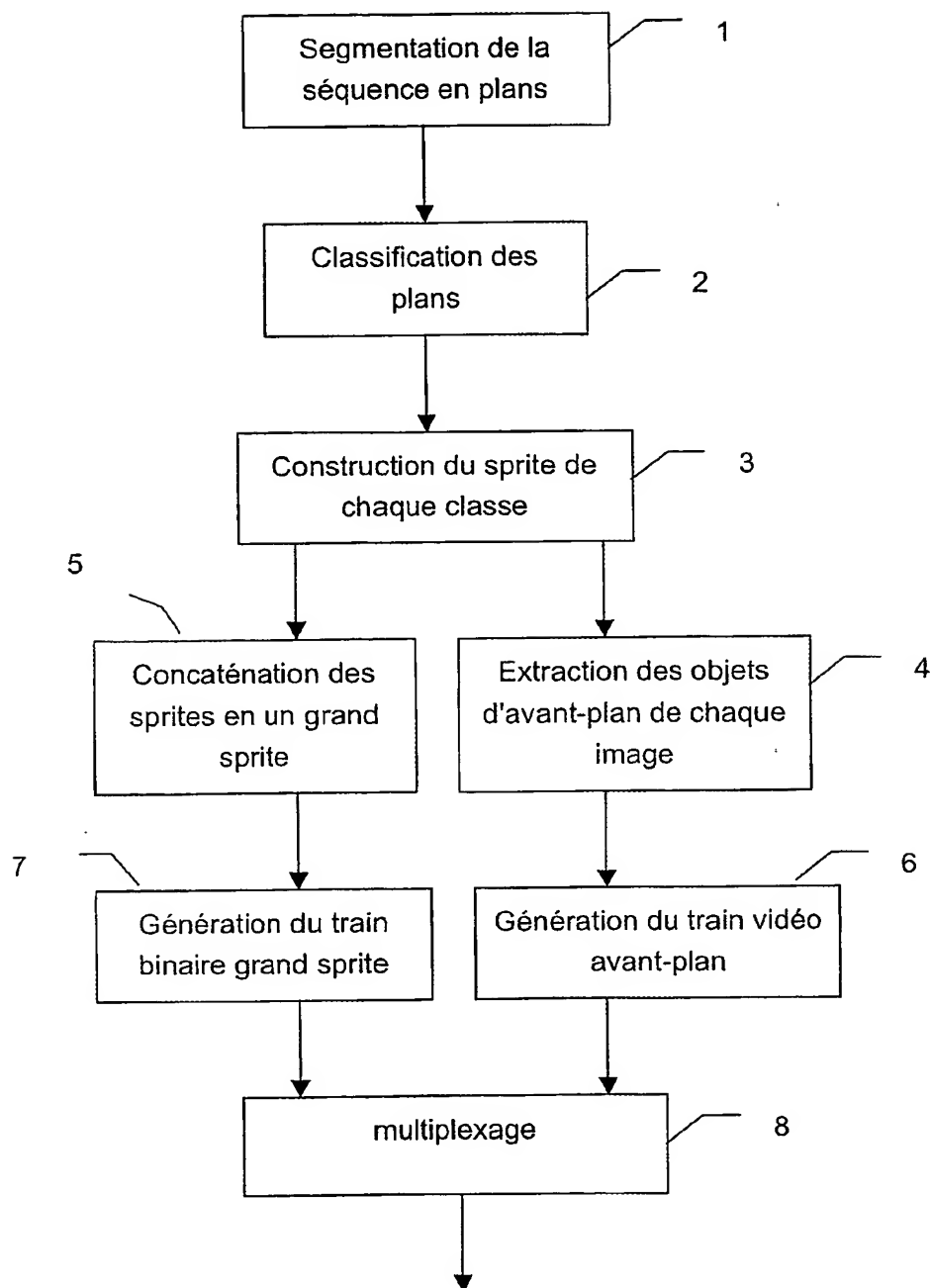


FIG.1

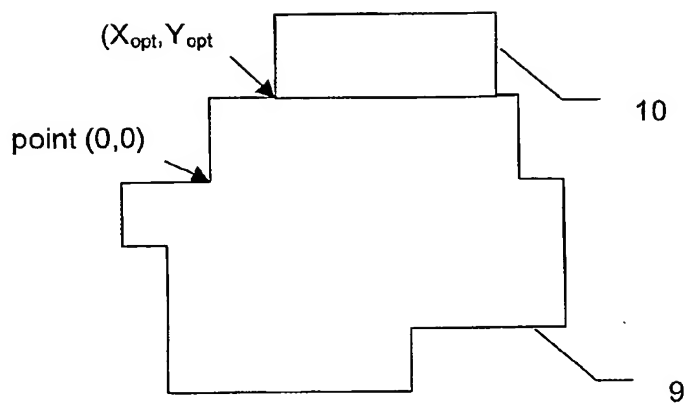


FIG.2

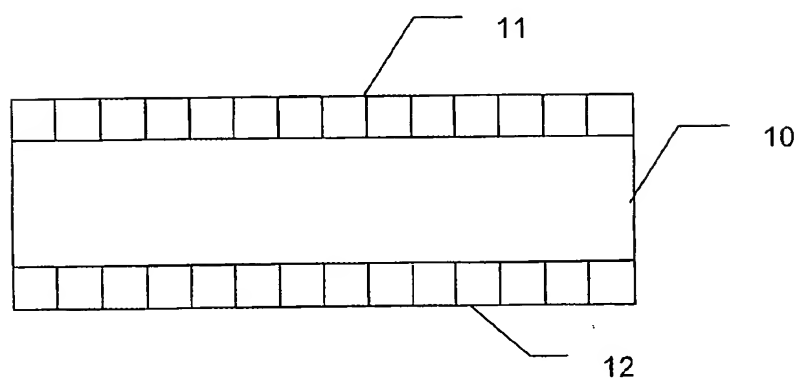


FIG.3

bloc B	bloc C
bloc A	bloc courant

FIG.4



# INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 03/50331

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC 7 H04N7/26		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) IPC 7 H04N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, PAJ, INSPEC, IBM-TDB		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GRAMMALIDIS N ET AL: "Sprite generation and coding in multiview image sequences" IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, MARCH 2000, IEEE, USA, vol. 10, no. 2, pages 302-311, XP002242024 ISSN: 1051-8215 page 302 -page 303, paragraph I page 304 -page 309, paragraphs III-V figures 4,7	1-8
Y	WO 98 02844 A (SARNOFF CORP) 22 January 1998 (1998-01-22) page 11, line 11 -page 16, line 31 --- -/--	1-8
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C.		
<input checked="" type="checkbox"/> Patent family members are listed in annex.		
<b>* Special categories of cited documents :</b>		
*A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *8* document member of the same patent family		
Date of the actual completion of the international search 12 December 2003		Date of mailing of the international search report 07/01/2004
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016		Authorized officer Lombardi, G

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 03/50331

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 00 08858 A (KONINKL PHILIPS ELECTRONICS NV) 17 February 2000 (2000-02-17) page 1, line 24 - line 25 ---	1-8
A	OHM J -R ET AL: "Incomplete 3D for multiview representation and synthesis of video objects" MULTIMEDIA APPLICATIONS, SERVICES AND TECHNIQUES - ECMAS'T'98. THIRD EUROPEAN CONFERENCE. PROCEEDINGS, MULTIMEDIA APPLICATIONS, SERVICES AND TECHNIQUES - ECMAS'T '98 THIRD EUROPEAN CONFERENCE PROCEEDINGS, BERLIN, GERMANY, 26-28 MAY 1998, pages 26-41, XP002242025 1998, Berlin, Germany, Springer-Verlag, Germany ISBN: 3-540-64594-2 page 5 -page 10, paragraph 3 figures 3-7 -----	1-8

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Publication No

PCT/EP 03/50331

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9802844 A	22-01-1998	CA 2261128 A1	22-01-1998
		EP 0979487 A1	16-02-2000
		JP 2002514359 T	14-05-2002
		KR 2000023784 A	25-04-2000
		WO 9802844 A1	22-01-1998
		US 6075905 A	13-06-2000
WO 0008858 A	17-02-2000	CN 1287749 T	14-03-2001
		CN 1287750 T	14-03-2001
		CN 1287751 T	14-03-2001
		WO 0008858 A1	17-02-2000
		WO 0008859 A1	17-02-2000
		WO 0008860 A1	17-02-2000
		EP 1042918 A1	11-10-2000
		EP 1042919 A1	11-10-2000
		EP 1042920 A1	11-10-2000
		JP 2002522982 T	23-07-2002
		JP 2002522836 T	23-07-2002
		JP 2002522983 T	23-07-2002
		US 6526183 B1	25-02-2003
		US 6496606 B1	17-12-2002
		US 6487323 B1	26-11-2002

# RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale No

PCT/EP 03/50331

**A. CLASSEMENT DE L'OBJET DE LA DEMANDE**  
CIB 7 H04N7/26

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

**B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE**

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 7 H04N

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

EPO-Internal, WPI Data, PAJ, INSPEC, IBM-TDB

**C. DOCUMENTS CONSIDERES COMME PERTINENTS**

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	GRAMMALIDIS N ET AL: "Sprite generation and coding in multiview image sequences" IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, MARCH 2000, IEEE, USA, vol. 10, no. 2, pages 302-311, XP002242024 ISSN: 1051-8215 page 302 -page 303, alinéa I page 304 -page 309, alinéas III-V figures 4,7	1-8
Y	WO 98 02844 A (SARNOFF CORP) 22 janvier 1998 (1998-01-22) page 11, ligne 11 -page 16, ligne 31 -/-	1-8

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

\* Catégories spéciales de documents cités:

- \*A\* document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- \*E\* document antérieur, mais publié à la date de dépôt international ou après cette date
- \*L\* document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- \*O\* document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- \*P\* document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

\*T\* document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

\*X\* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

\*Y\* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

\*Z\* document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

12 décembre 2003

Date d'expédition du présent rapport de recherche internationale

07/01/2004

Nom et adresse postale de l'administration chargée de la recherche internationale  
Office Européen des Brevets, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel (+31-70) 340-2040, Tx. 31 651 epo nl  
Fax (+31-70) 340-3016

Fonctionnaire autorisé

Lombardi, G

# RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale No

PCT/EP 03/50331

## C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	WO 00 08858 A (KONINKL PHILIPS ELECTRONICS NV) 17 février 2000 (2000-02-17) page 1, ligne 24 - ligne 25 -----	1-8
A	OHM J -R ET AL: "Incomplete 3D for multiview representation and synthesis of video objects" MULTIMEDIA APPLICATIONS, SERVICES AND TECHNIQUES - ECAST'98. THIRD EUROPEAN CONFERENCE. PROCEEDINGS, MULTIMEDIA APPLICATIONS, SERVICES AND TECHNIQUES - ECAST '98 THIRD EUROPEAN CONFERENCE PROCEEDINGS, BERLIN, GERMANY, 26-28 MAY 1998, pages 26-41, XP002242025 1998, Berlin, Germany, Springer-Verlag, Germany ISBN: 3-540-64594-2 page 5 -page 10, alinéa 3 figures 3-7 -----	1-8

# RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale No

PCT/EP 03/50331

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9802844 A	22-01-1998	CA 2261128 A1	22-01-1998
		EP 0979487 A1	16-02-2000
		JP 2002514359 T	14-05-2002
		KR 2000023784 A	25-04-2000
		WO 9802844 A1	22-01-1998
		US 6075905 A	13-06-2000
WO 0008858 A	17-02-2000	CN 1287749 T	14-03-2001
		CN 1287750 T	14-03-2001
		CN 1287751 T	14-03-2001
		WO 0008858 A1	17-02-2000
		WO 0008859 A1	17-02-2000
		WO 0008860 A1	17-02-2000
		EP 1042918 A1	11-10-2000
		EP 1042919 A1	11-10-2000
		EP 1042920 A1	11-10-2000
		JP 2002522982 T	23-07-2002
		JP 2002522836 T	23-07-2002
		JP 2002522983 T	23-07-2002
		US 6526183 B1	25-02-2003
		US 6496606 B1	17-12-2002
		US 6487323 B1	26-11-2002

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**